

Abstract of the Invention

[0050] A document digitizing method digitizes and automatically indexes documents in printed form. The method includes optically scanning the document, forming and storing a digitized image file from the optically scanned document, optically recognizing characters in the optically scanned document, and forming and storing a text file of the optically recognized characters in document. A retrieval method for retrieving the digitized image file for a document includes searching the text files to identify any having a selected text string and providing access to the digitized image files that correspond to those text files. The digital image file and the text file together represent a digitized document data structure that combines a digital image of a document with a text file of optically recognized characters in the digital image.